图书情报工作
第 65 卷 第 8 期 2021 年 4 月

网络突发事件中社交机器人情感的交互式影响机制研究*

■ 马晓悦[1] 孟啸[1] 王镇[1] 刘益东[2]

[1] 西安交通大学新闻与新媒体学院 西安 710049 [2] 上海市浦东微热点大数据研究院 上海 201203

摘 要：[目的/意义]细化社交机器人对网民情感的干预机制，同时从实践上为网络突发事件舆情治理提供建议参考。[方法/过程]以仁济医院赵晓菁事件为例，利用朴素贝叶斯方法计算微博情感倾向，通过构建向量自回归模型(VAR)并进行格兰杰因果检验、脉冲响应分析以及方差分解分析，确定社交机器人、意见领袖与普通用户在事件生命周期各阶段的情感关系。[结果/结论]社交机器人、意见领袖与普通用户的情感关系随舆情阶段演进发生变化，在爆发期，社交机器人放大了意见领袖对普通用户的情感影响；在成熟期，社交机器人影响式微，普通用户的情感反作用于社交机器人与意见领袖；在衰退期，三者保持较为独立的情感关系。此外，社交机器人的影响策略具有隐匿性和间接性特征。

关键词：突发事件 社交机器人 情感分析 时间序列分析

分类号：G250

DOI：10.13266/j.issn.0252-3116.2021.08.008

1 引言

近年来，社交媒体的普及为突发事件提供了迅速扩散的渠道，而越来越多的研究发现，社交机器人发布的信息占据了社交媒体信息相当大的比重[1]。社交机器人的网络参与极易干扰公众对事件的认知和判断，催生网民群体的负面情绪，增强公众网络讨论的情感极性[2]，这大大增加了网络突发事件舆情的不确定性。因此，从情感的角度出发，研究社交机器人与人类用户在突发事件网络舆情演化过程中的情感交互影响具有重要的现实治理意义。

社交机器人由算法驱动，能够在社交媒体中模拟人类的信息行为。已有许多研究指出了社交机器人的应用风险，包括在政治选举中传播假新闻[3]，在自然灾害等紧急情况下传播错误和虚假信息[4]。在 COVID-19 爆发期间就有研究发现 Twitter 上的社交机器人大量转发来自低可信度信源的信息以及具有政治偏见的内容[5]。然而，以网络突发事件为背景的社交机器人研究仍在初始阶段。首先，以往对社交机器人的研究多是基于 Twitter 等国外社交媒体，缺少中文社交媒体环境下的研究。其次，情感是推动网络舆情产生、形成和演变的根本力量和深层逻辑[6]，社交机器人与人类情感的交互特征反映了人机关系的潜在内核，过去的研究重视对社交机器人内容与行为特征的静态描述，较少验证社交机器人与人类情感在统计意义上的交互影响。此外，传统的网络突发事件研究多以政府部门、网络媒体、普通网民、网络水军等作为舆情主体展开研究，缺少对社交机器人的关注。

鉴于此，本研究试图把社交机器人作为新的参与主体嵌入到突发事件的研究框架之中，揭示社交机器人与意见领袖、普通用户在网络突发事件情境下的情感交互关系，特别是这种交互关系在事件不同生命周期阶段的差异性。本文在研究中主要解决以下问题：①社交机器人与意见领袖、普通用户的情感存在何种关系，这种关系在网络突发事件各阶段如何变化？②社交机器人通过何种行为策略对人类情感产生影响？

为解决上述问题，本研究基于实际突发事件微博

* 本文系教育部人文社会科学研究规划基金“信息协同视角下基于可视化媒介的智慧应急响应行为研究”(项目编号：19YJA870009)和陕西省自然科学基础研究计划一般项目－面上项目“基于散射－叠加效应的新媒体信息演化模型构建及事件类别判定研究”(项目编号：2020JM－056)研究成果之一。

作者简介：马晓悦(ORCID：0000－0003－4932－6450)，特聘研究员，博士，博士生导师，E-mail：xyma_mail@163.com；孟啸(ORCID：0000－0002－7551－5856)，硕士研究生；王镇(ORCID：0000－0002－1859－7605)，硕士研究生；刘益东(ORCID：0000－0002－0771－6365)，学士。

收稿日期：2020－10－16 修回日期：2021－01－11 本文起止页码：74－84 本文责任编辑：易飞

马晓悦，孟啸，王镇，等.网络突发事件中社交机器人情感的交互式影响机制研究[J].图书情报工作,2021,65(8):74－84.

数据，在划分用户类别与事件生命周期的基础上，提取社交机器人与意见领袖、普通网民的情感倾向值，利用向量自回归模型和广义脉冲响应函数分析等方法分析社交机器人、意见领袖与普通用户微博的情感关系。

2 文献回顾

2.1 社交机器人

社交机器人(social bots)是社交媒体中自动生成内容并和人类互动，试图模仿并可能会改变人类的计算机算法[7]。一些社交机器人可以提供服务，如在自然灾害和危机中进行自动报警[8]，也有部分社交机器人被用作“网络水军”，具有社会危害性。现有对社交机器人的研究主要集中于两个方面：①社交机器人识别监测；②社交机器人的活动及其影响。

在社交机器人识别方面，依赖社会网络特征来区分真实用户和社交机器人的方法得到较为广泛的使用。Y. Boshmaf 等提出了一个名为 Integro 的自动虚假账户检测系统，利用有监督机器学习和用户排名方案将大多数真实账户排名高于虚假账户[9]。S. Hurtado 等假设具有高度关联的类似时间活动的账户很可能是机器人，他们利用时间和网络信息在 Reddit 平台上检测政治机器人[10]。另一种常见的识别方法是基于机器学习的，著名的 BotOrNot 系统就是通过一千多个行为特征来训练分类器，进而衡量一个账号是社交机器人的可能性[11]。除行为以外，文本信息和内容分析也被尝试用于社交机器人识别[12]。

在社交机器人行为特征及其影响层面，研究者广泛关注社交机器人在政治活动中的表现。社交机器人通常被认为会对社交媒体舆论进行操纵、改变公众对政治实体的认知、甚至会对政治选举结果产生影响[13]。例如，在 2016 年美国大选期间，一些候选人使用自动账户或机器人来提高自己的社交媒体关注度和粉丝数[14]。在英国脱欧中，大量社交机器人只在公投辩论期间发推特，投票后就消失[15]。关于 Twitter 上的中国议题，有研究认为社交机器人的存在可增加人类用户对于特定信息的接触，改变既有的信息交互结构[16]。张洪忠总结了政治机器人的 5 种策略：营造虚假人气，推送大量政治消息，传播虚假或垃圾政治信息，制造烟雾遮蔽效应混淆公众视听，塑造高度人格化形象的虚拟意见领袖[17]。但也有研究认为社交机器人的影响并不显著，F. Brachten 等发现在 2017 年德国州选举期间，已识别的社交机器人并没有集体政治策略的迹象，影响力也微乎其微[18]。除了上述政治议题外，公共卫生领域中的反疫苗运动[19]和电子烟推广[20]等话题中也发现了社交机器人的参与。

2.2 网络突发事件中的情感分析

情感分析是涉及分析人们对不同实体的情感、态度和观点的研究领域。网络突发事件中的情感分析对了解网民观点和制定决策具有重要意义，相关研究主要集中于信息监测、情感演化模式、情绪扩散模式和政府互动策略等方面。从信息监测层面，有学者建立基于地理位置的突发事件主题、情感监测模型来反映情感在地区间的动态转移[21]。在情感演化层面，毛太田等通过对上海警察绊摔小孩事件微博的分析发现在信息不明朗情况下，公众易被舆论场中最显著的意见所影响[22]。从情绪扩散的角度，S. Stieglitz 等基于 Twitter 上超过 16.5 万推文的研究表明情绪激动的推文比中性的推文更容易得到转发[23]。从政府信息行为与公众情绪的关系角度，W. Zhang 等的研究发现政府信息发布策略对焦虑、厌恶情绪的蔓延有显著负面影响[24]。

以往研究者使用的研究方法主要包括基于情感词典的情感计算和基于神经网络、深度学习等机器学习技术的分类算法[25]。前者主要是通过对意见词的情感标注来决定词的情感倾向，主要有 3 种建立情感词典的方法，分别是手动、语料库和基于字典的方法，手动法相对于后两者更为费时，它通常与其他两者结合来避免自动化技术发生的错误[26]。F. Aslam 等利用 NRC 情感词典计算新冠疫情期间新闻头条在 8 个维度的情感表现[27]。国内学者安璐等以塞卡事件为例，参考大连理工大学情感词典并利用点互信息的方法构建表情符号词典来计算微博文本情感[28]。张鹏等结合扎根理论的质性研究特点构建突发事件专属情感词典，一定程度上提高了情感分析的准确性[29]。机器学习方法使用完整数据的一部分来训练分类器，不依赖于先前的词汇，典型的方法有朴素贝叶斯、最大熵和支持向量机等。例如，邓君等通过 Word2Vec 和 SVM 方法构建情感分类模型划分滴滴温州女孩遇害话题评论内容的情感极性[30]。吴鹏等使用具有情感意义的词向量作为文本特征，以双向长短期记忆模型作为分类器，对网民情感极性进行分类[31]。虽然机器学习方法被广泛应用，但它需要大量的训练数据并对训练数据的质量有高要求，而且还存在黑箱和难以解释等问题。

综上所述，以往研究已经在社交机器人识别与情感分析技术方面进行了大量尝试，但仍存在一些不足：①虽然先前对社交机器人影响的研究更多强调其对网

络舆论的操纵作用，但这种影响较少被证实，更多研究侧重于对社交机器人内容与行为特征的描述，并且案例多来自于国外社交媒体与政治领域话题。②已有的网络突发事件情感分析主要涉及政府、媒体、意见领袖、网民等主体，缺少对社交机器人的关注，尚未有研究将社交机器人、意见领袖与普通用户的情感作为 3 个变量放入同一时间序列系统中进行分析。因此，有必要在网络突发事件的情境下检验社交机器人与意见领袖、普通用户的情感关系，为舆情监测与处置提供参考。

3　研究方法

本研究以具体网络突发事件为案例，采用向量自回归模型分析社交机器人、意见领袖与普通用户情感的交互影响。具体研究过程可分为数据获取与预处理、情感倾向提取、情感关系分析 3 个部分。首先，对获取到的社交机器人与人类微博数据进行预处理，去除其中的无意义信息并划分参与主体与事件生命周期；然后，计算微博的情感倾向并获得社交机器人与意见领袖、普通用户的情感变化曲线；最后，对事件各阶段的社交机器人与意见领袖、普通用户的情感时间序列建立向量自回归模型，分析三类参与主体的情感交互关系。

3.1　数据获取与预处理

3.1.1　数据获取与文本清洗

本研究选取“仁济医院赵晓菁事件”作为突发事件案例。该事件发生在 2019 年 4 月 24 日，上海仁济医院专家赵晓菁因拒绝接诊插队病人并与病人发生争执而被警方戴上手铐带走，在 4 月 26 日该事件在新浪微博平台上被曝光，引发社会广泛关注。本研究所使用的微博数据集由新浪微热点大数据研究院合作提供，微热点大数据研究院（www.wrd.cn）是新浪微博旗下的大数据传播应用平台，提供新媒体大数据服务。该数据集经过脱敏处理，包括事件发生前后与“仁济医院赵晓菁事件”相关的 351 257 条微博，并且提供了发布者是否为社交机器人的标注，每条数据包含了转发微博内容、微博发布时间、原微博内容、原微博发布时间、认证类型以及社交机器人身份标注字段，由于知识产权保护问题，该数据集尚不能公开获取。在获取到初始微博数据后，我们选取了事件主要讨论期“2019－4－26”至“2019－5－2”时间段内的相关微博，共 346 447 条。利用正则表达式剔除微博文本中“@用户名”“转发微博”等干扰情感分析结果的字段。

3.1.2　意见领袖与普通用户划分

在两级传播理论中，意见领袖是信息传播中的重要中介[32]，他们往往拥有较高的社会地位和广泛的社会网络，能够影响和塑造他人观点[33]。为区分意见领袖与社交机器人对普通用户情感的影响，本研究将人类用户细分为意见领袖与普通用户两类。微博中的认证用户往往是各行业的专家、名人，是否认证对意见领袖影响力具有重要作用[34]。为提取事件中的意见领袖，本文基于认证类型对人类用户进行分类，将认证类型为“金 V”“橙 V”“蓝 V”“达人”的用户划分为意见领袖，把其他人类用户划分为普通用户。

3.1.3　生命周期划分

在以往研究中，学者根据其研究目的对网络舆情生命周期进行了不同的划分。如王国华等将药家鑫事件舆情划分为发生、发展、演化、消解 4 个阶段[35]；李纲等将突发事件网络舆情划分为潜伏、爆发、成熟和衰退 4 个阶段[36]；谢科范等将网络舆情突发事件分为潜伏期、萌动期、加速期、成熟期、衰退期 5 个阶段[37]。结合以往研究使用的划分方法，考虑到本文所选取的事件舆情生命周期短，发生发展迅速，依据微博讨论数量曲线把“仁济医院赵晓菁事件”划分为爆发、成熟和衰退 3 个主要阶段。

3.2　社交机器人、意见领袖与普通用户的情感倾向分析

SnowNLP 是受 TextBlob 启发而编写的中文自然语言处理库，它提供了一种基于朴素贝叶斯算法的文本情感倾向计算方法，由于初始的情感分析功能以购物评价作为训练数据，因此本文使用带有人工情感倾向标注的微博语料训练情感倾向判别模型，使其适用于微博情感分析任务，在测试数据集上实现 90.5% 的情感分类准确率。随后使用该模型判断情感倾向，得到单条微博文本情感为正向的概率，概率值的取值范围在 0 到 1 之间，越接近 1 则表明文本情感越正面，越接近 0 则表示情感越负面。然后，考虑到突发事件舆情周期的短暂性，以 10 分钟为单位对微博数据进行切片，计算每个时间切片上的平均情感值，最终得到意见领袖、社交机器人和普通用户的情感值时间序列数据与情感演化曲线。

3.3　社交机器人、意见领袖与普通用户的情感关系分析

社交机器人与意见领袖、普通用户的情感处在动态交互的系统中，传统的多元线性回归难以对其进行解释，因而本研究采用经济学中常用的向量自回归模型（vector dutoregression model, VAR）分析社交机器人与意见领袖、普通用户间的情感关系。VAR 模型把系

统中每一个内生变量作为系统中所有内生变量的滞后值的函数来构造模型，从而将单变量自回归模型推广到多元时间序列变量组成的向量自回归模型。在本研究中，我们针对舆情 3 个主要阶段（爆发期、成熟期、衰退期）分别建立 VAR 模型，并基于 VAR 模型进行格兰杰因果关系分析、广义脉冲响应函数分析和方差分解分析。

3.3.1 VAR 模型构建

当时间序列数据不稳定时，可能导致伪回归而影响结果可靠性。因此，在利用 VAR 模型进行研究前，我们首先对各生命周期上的数据进行 ADF 单位根检验，P 值均显著，表明各组变量数据都是平稳的，可以直接进行格兰杰因果关系检验和建立 VAR 模型。

VAR 模型的建立需要确定最优滞后阶数来保证滞后期与自由度间的平衡。本研究综合考虑 LR、FPE、AIC、SC 和 HQ 等检验标准，确定爆发期、成熟期、衰退期三个模型的最优滞后阶数分别为滞后 1 期、滞后 1 期和滞后 2 期。然后，依据最优滞后期分别建立 VAR 模型，对应的 VAR(p) 模型形式如下：

$$y_t = A_0 + A_1 y_{t-1} + \cdots + A_p y_{t-p} + e_t \qquad \text{公式(1)}$$

其中，$y_t = \begin{pmatrix} OpinionLeader \\ Crowd \\ SocialBots \end{pmatrix}_t$，OpinionLeader、Crowd 和 SocialBots 分别是意见领袖、普通用户、社交机器人的情感倾向，A_0 是常数向量，e_t 是随机扰动项矩阵，A_t 是系数矩阵，p 是滞后阶数。

为了确定模型的平稳性和结果的正确性，需要对模型进行稳健性检验。AR 根检验的结果显示特征根模的倒数均小于 1，表明模型稳定，无需重新构建。

3.3.2 格兰杰因果关系检验

为了判断社交机器人、意见领袖、普通用户情感的影响与被影响关系，我们应用格兰杰因果关系检验。格兰杰因果关系检验的基本思想是若一时间序列变量 x 是另一变量 y 的格兰杰原因，则意味着在预测变量 x 的模型中加入 y 的过去值的预测结果比仅仅使用 x 的过去值的预测结果更准确。在本文的 VAR(p) 模型中，若检验主体 A 的情感 y_j 是否为主体 B 的情感 y_i 的格兰杰原因，则格兰杰因果关系检验的原假设为主体 A 的情感不是主体 B 的情感的格兰杰原因，即 y_j 的参数 $a_{ij}^{(q)}$，$q=1,2,3,\cdots,p$ 整体为 0，格兰杰检验的结果通过 F 检验来完成。

3.3.3 广义脉冲响应函数分析

脉冲响应函数描述的是某一内生变量随机误差项施加一个标准差大小的冲击后对另一内生变量的当期值和未来值产生的影响。下面以 VAR(2) 模型为例子来说明脉冲响应函数的基本思路。

$$\begin{cases} x_t = a_1 x_{t-1} + a_2 x_{t-2} + b_1 z_{t-1} + b_2 z_{t-2} + \varepsilon_{1t} \\ z_t = c_1 x_{t-1} + c_2 x_{t-2} + d_1 z_{t-1} + d_2 z_{t-2} + \varepsilon_{2t} \end{cases} \qquad \text{公式(2)}$$

假定上述系统从 $t=0$ 开始活动，设 $x_1 = x_2 = z_1 = z_2 = 0$，于第 0 期设定扰动项 $\varepsilon_{10}=1$，$\varepsilon_{20}=0$，其后它们均为 0。那么初期给予的扰动将在系统中不断传递，通过迭代计算可以得到 $x_0, x_1, x_2, x_3, \cdots$，称为由 x 的脉冲引起的 x 的响应函数，同理可求得 $z_0, z_1, z_2, z_3, \cdots$，称为由 x 的脉冲引起的 z 的响应函数。当第 0 期的脉冲为 $\varepsilon_{10}=0$，$\varepsilon_{20}=1$ 时，可以求得 z 的脉冲引起的 x 和 z 的响应函数。本文运用 H. H. Pesaran 和 Y. Shin 提出的广义脉冲响应函数法（GIRF）来进行分析，该方法的结果不受变量排序关系的影响[38]。在本研究中，广义脉冲响应函数分析的结果反映了意见领袖、普通用户、社交机器人中一类主体情感的波动变化引起的另一类主体情感变化的大小、方向与滞后特征。

3.3.4 方差分解分析

方差分解能够刻画结构冲击对所研究内生变量变化水平的贡献程度，并借此进而评价不同结构冲击的重要性问题。由方程（1）得到的向量移动平均模型（Vector Moving Average Model，VMA）为：

$$y_{it} = \sum_{j=1}^{k} \left(c_{ij}^{(0)} \varepsilon_{jt} + c_{ij}^{(1)} \varepsilon_{jt-1} + c_{ij}^{(2)} \varepsilon_{jt-2} + \cdots\right) \qquad \text{公式(3)}$$

其中，$c_{ij}^{(q)}$ 为 y_j 的脉冲在第 q 期引起的 y_i 的响应函数，据此本文采用的方差分解模型为：

$$RVC_{j\to i}(\infty) = \frac{\sum_{q=0}^{\infty} (c_{ij}^{(q)})^2 \sigma_{jj}}{VAR(y_{it})} = \frac{\sum_{q=0}^{\infty} (c_{ij}^{(q)})^2 \sigma_{jj}}{\sum_{j=1}^{k} \{\sum_{q=0}^{\infty} (c_{ij}^{(q)})^2 \sigma_{jj}\}} \qquad \text{公式(4)}$$

$RVC_{j\to i}(\infty)$ 是相对方差贡献度，取值在 0 到 1 之间，即根据第 j 个变量基于冲击的方差对 y_i 的方差的相对贡献度来观测第 j 个变量对第 i 个变量的影响。$\sum_{q=0}^{\infty}(c_{ij}^{(q)})^2 \sigma_{jj}$ 是第 j 个扰动项对第 i 个变量从无限过去到现在时点的影响用方差加以评价的结果。实际上，如果模型满足平稳性条件，则 $c_{ij}^{(q)}$ 随着 q 的增大呈现几何级数性的衰减，所以只需取有限的 s 项即可[39]。在本研究中，我们基于方差分解来识别意见领袖、普通用户、社交机器人之间影响的贡献程度问题。

4 研究结果

4.1 数据描述

本文以“仁济医院赵晓菁事件”为例，采取 3.2 节

所述的方法绘制微博舆情演化趋势图并将事件舆情生命周期划分为爆发期、成熟期和衰退期 3 个阶段，如图 1 所示。仁济医院赵晓菁事件微博传播呈现“单峰”特征，在图示的阶段划分中，4 月 26 日达到了微博发布数量的最高峰，是爆发期的显著特征；随后在 4 月 27 日和 28 日，事件热度下降，微博数量出现数个小高峰；到 29 日，微博发布数减少到较低水平且变化平缓，事件舆情进入衰退期。各主体的微博发布数量状况见表 1，社交机器人主要在爆发期和成熟期发布微博，在此次事件中共发布了 29 031 条微博，占总微博数的 8.4%；意见领袖和普通用户的微博发布数量分别占比 15.2% 和 76.4%。

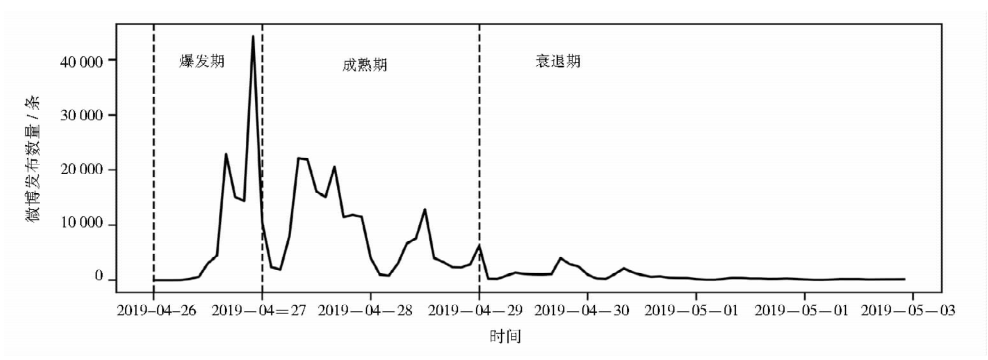


图 1 “仁济医院赵晓菁事件”生命周期

表 1 “仁济医院赵晓菁事件”各阶段的微博发布数量

（单位/条）

主体	爆发期	成熟期	衰退期	总计
意见领袖	16 132	30 880	5 737	52 749
普通用户	76 985	159 987	27 695	264 667
社交机器人	12 253	13 394	3 384	29 031
总计	105 370	204 261	36 816	346 447

本研究进而对三类主体的行为与内容特征进行描述，各类主体的微博发布频率、来源以及内容特征见表 2。在微博发布频率方面，社交机器人微博发布频率较高，平均每个社交机器人发布 2.460 条微博。在微博来源方面，社交机器人、意见领袖与普通用户主要是通过转发他人微博参与网络突发事件讨论，意见领袖是三类主体主要的微博来源，社交机器人的微博较少被转发；相较而言，意见领袖微博原创率较高，为 7.7%，社交机器人微博原创率较低。在内容特征方面，社交机器人发布微博的内容较短，平均为 18.647 个字符，表情符号较多，平均每条微博包含 1.344 个表情符号，平均话题标签数与提到用户数较少，表明其较少在话题下进行讨论，较少与其他用户进行互动。

表 2 社交机器人、意见领袖、普通用户的行为与内容特征

主体	行为特征					内容特征			
	微博发布频率	微博来源				平均内容长度	平均表情符号数	平均话题标签数	平均提到用户数
		社交机器人	意见领袖	普通用户	原创				
社交机器人	2.460	0.11%	93.39%	2.60%	3.90%	18.647	1.344	0.090	0.154
意见领袖	1.610	0.29%	88.19%	3.78%	7.74%	58.625	0.513	0.155	1.114
普通用户	1.542	0.14%	91.53%	4.15%	4.18%	39.993	0.532	0.233	0.798

4.2 社交机器人、意见领袖、普通用户的情感变化

为观察社交机器人情感的变动特征，以 3 小时为单位的微博情感变化曲线见图 2，两条纵向虚线区分了爆发期、成熟期和衰退期 3 个舆情阶段。从整体来看，三者的微博情感倾向较为积极，社交机器人情感值相较人类变动更剧烈，意见领袖与普通用户的情感较为接近。比较各舆情阶段的情感值曲线可以发现，在爆发期和成熟期，社交机器人与意见领袖、普通用户情感保持较为一致的变化趋势；衰退期的曲线波动更加剧烈，社交机器人与人类情感差异较大。情感倾向时间序列为接下来建立 VAR 模型分析情感交互关系提供了数据支撑。

4.3 社交机器人、意见领袖与普通用户的情感关系

4.3.1 社交机器人、意见领袖与普通用户情感的格兰杰因果关系

基于 VAR 模型的格兰杰因果关系检验的结果见表 3，它反映了各阶段社交机器人、意见领袖与普通用户情感的预测关系。在舆情爆发期，在 5% 的显著性

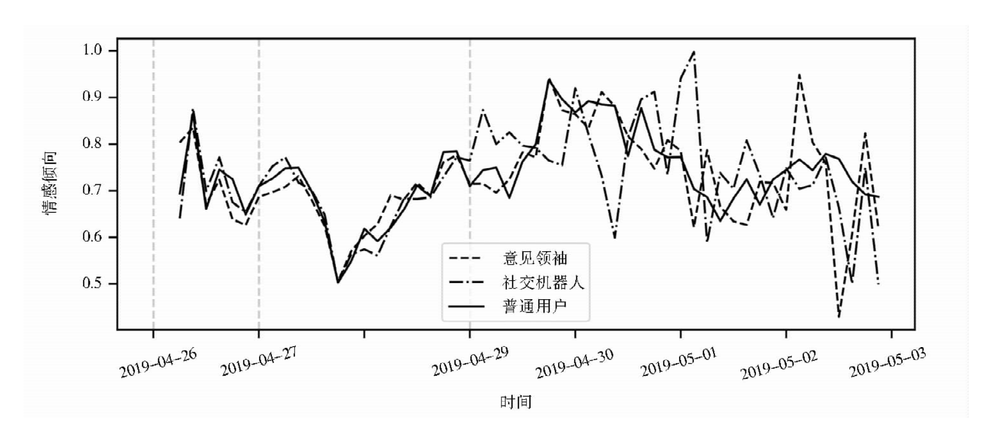


图2　意见领袖、普通用户、社交机器人情感演化曲线

水平上，意见领袖和社交机器人是普通用户情感的格兰杰原因，这说明意见领袖和社交机器人的情感变化领先于普通用户；在成熟期，意见领袖是普通用户的格兰杰原因，普通用户是社交机器人与意见领袖情感的格兰杰原因；在衰退期，意见领袖是普通用户与社交机器人情感的格兰杰原因，普通用户是意见领袖情感的格兰杰原因。这说明在舆情演化过程中，社交机器人、意见领袖与普通用户的情感变化具有一定联系。

表3　格兰杰因果关系检验结果

原假设	爆发期			成熟期			衰退期		
	Chi-sq	Prob.	结论	Chi-sq	Prob.	结论	Chi-sq	Prob.	结论
社交机器人不是意见领袖情感的格兰杰原因	3.481	0.176	接受	2.516	0.113	接受	1.892	0.388	接受
普通用户不是意见领袖情感的格兰杰原因	4.617	0.099	接受	39.124	0.000	拒绝	16.002	0.000	拒绝
意见领袖不是社交机器人情感的格兰杰原因	4.999	0.082	接受	0.062	0.803	接受	6.823	0.033	拒绝
普通用户不是社交机器人情感的格兰杰原因	1.159	0.560	接受	39.517	0.000	拒绝	0.206	0.902	接受
社交机器人不是普通用户情感的格兰杰原因	8.091	0.018	拒绝	0.024	0.878	接受	0.197	0.906	接受
意见领袖不是普通用户情感的格兰杰原因	8.557	0.014	拒绝	25.392	0.000	拒绝	6.62	0.037	拒绝

4.3.2　社交机器人、意见领袖与普通用户情感的动态影响

意见领袖、社交机器人以及普通用户自身对普通用户情感的影响见图3－1、3－2和3－3，社交机器人于意见领袖的情感影响具有一定时滞。在爆发期，如图3－1所示，当在本期给意见领袖和社交机器人情感一个正冲击，普通用户情感在第2期和第3期分别上升了约4.32%和3.70%。社交机器人情感的正向冲击也对人群情绪产生了积极的影响，使第2期普通用户情感上升了3.52%，第3期上升3.65%。但在接下来的两个舆情阶段（见图3－2和图3－3），社交机器人和意见领袖的影响力削弱，社交机器人的影响趋近于0。

普通用户和社交机器人情感的正冲击在事件3个阶段上对意见领袖情感的影响见图3－4、3－5和3－6。可以看到，在爆发期，普通用户和社交机器人对意见领袖情感的影响分别在当期和第3期达到最大；在成熟期，普通用户和社交机器人对意见领袖情感的影响在第2期达到最大，社交机器人影响衰退较为明显。到了衰退期，社交机器人对意见领袖情感的影响在当期达到最大后迅速削弱。

普通用户和意见领袖情感的正冲击在事件3个阶段上对社交机器人的影响见图3－7、3－8、3－9。在爆发期，普通用户和和意见领袖对社交机器人情感的影响在当期达到最大后收敛；在成熟期，普通用户对社交机器人情感的影响在第2期达到最大，意见领袖对社交机器人情感的影响在第3期达到最大。在衰退期，意见领袖情感的正向冲击会在短期内造成社交机器人的情感波动，普通用户对社交机器人情感的影响趋近于0。

4.3.3　社交机器人、意见领袖与普通用户情感的影响重要性

意见领袖情感在爆发期、成熟期和衰退期的方差分解结果见图4－1、4－2、4－3。在爆发期，滞后1到10期对意见领袖情感贡献率最大的是其自身因素，但其他因素影响逐渐增加，社交机器人和普通用户对意见领袖的情感贡献率在第10期分别增加至16.55%和2.57%；在成熟期，如图4－2所示，社交机器人对意见

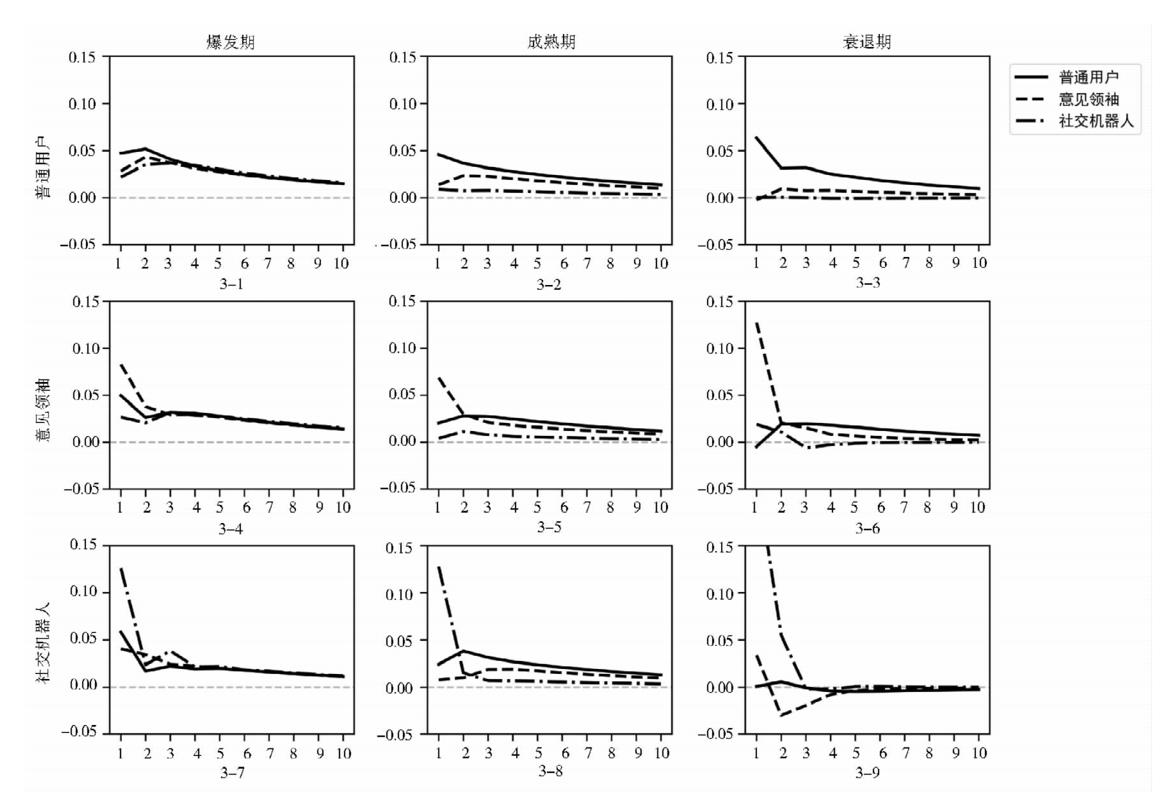


图 3　事件各阶段变量的冲击响应

领袖情感变动的贡献趋近于 0，普通用户对意见领袖情感变动的贡献率在第 10 期达到 23.32%；进入衰退期后，社交机器人与普通用户对意见领袖情感的贡献率较成熟期有所下降。

社交机器人情感在事件三个阶段的方差分解结果见图 4－4、4－5 和 4－6。在爆发期，意见领袖对社交机器人情感的贡献率维持在 20% 左右，普通用户对社交机器人情感贡献在 0 附近；在成熟期，普通用户对社交机器人情感的贡献增加，在第 10 期达到 16.4%，意见领袖对社交机器人情感的贡献最大值在第 10 期达到 8.59%，相较爆发期有所下降；进入衰退期后，意见领袖和普通用户的贡献率均在 5% 以下。

普通用户情感在事件 3 个阶段的方差分解结果见图 4－7、4－8、4－9。在爆发期，普通用户情感变动对其自身的贡献率由 56.12% 下降到 17.60%，意见领袖对普通用户情感的贡献率在第 2 期达到 50.47% 后趋于稳定，社交机器人对普通用户情感的贡献在第 10 期达到 27.10%；在成熟期，社交机器人对普通用户情感的贡献率维持在 3% 左右，意见领袖对普通用户情感的贡献率在第 10 期达到 33.49%；进入衰退期后，意见领袖和社交机器人对普通用户情感的影响趋近于 0，普通用户情感主要受其自身因素影响。

5　讨论

5.1　社交机器人、意见领袖与普通用户的情感关系具有阶段性差异

本研究发现随着事件的演化，社交机器人、意见领袖、普通用户的情感关系呈现不同特征：①在爆发期，意见领袖不仅对普通用户和社交机器人的情感具有较强的影响效果，还能够较大程度解释社交机器人的情感变化，同时，社交机器人情感的变动也能够预测和解释普通用户的情感变动；这表明社交机器人放大了意见领袖对普通用户情感的影响，普通用户处于相对被动的地位。在特定情况下，社交机器人的参与增加了情感极化的风险。②在成熟期，意见领袖和普通用户能够在较大程度上影响对方情感，社交机器人的情感受普通用户影响较大，对普通用户与意见领袖情感的影响效果微乎其微；这意味着在成熟期社交机器人的影响式微，意见领袖的引导作用相对削弱，普通用户的情感影响占据主导地位。③在衰退期，社交机器人与意见领袖、普通用户间仍存在预测关系，但三者情感的影响效果都较弱，这说明三者的情感关系此时较为独立。这与 L. Zhang 等的研究发现一致，情绪在突发事件发展初期更容易传播[40]。

关于社交机器人、意见领袖与普通用户的情感关

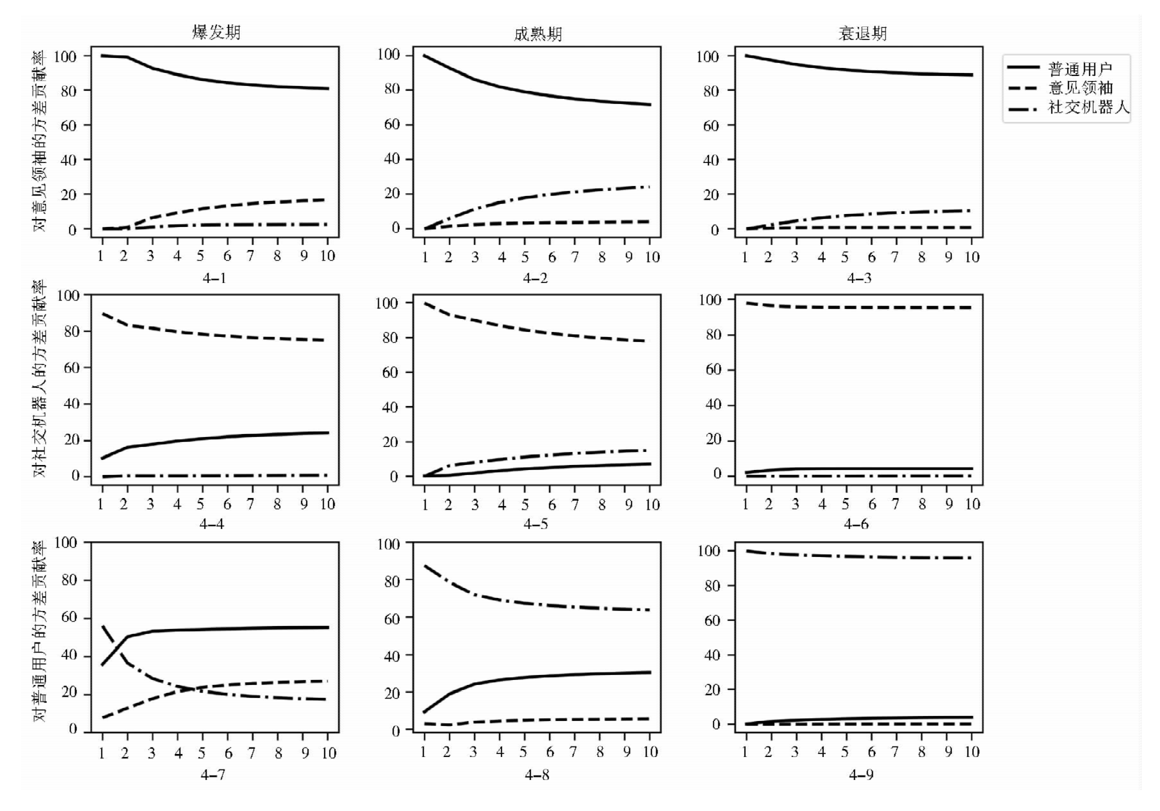


图 4　事件各阶段变量的方差贡献



系的演化机制，可从生命周期理论的视角进行解释：在爆发期，面对事件性质和相关信息的不确定性，普通用户的情感容易受到意见领袖的影响。进入成熟期后，公众对事件形成了相对稳定的态度和观点，意见领袖作为有影响力的群体，在发布和传播信息时会更加注重信息所产生的影响以及造成的后果，以维护自身的地位和社会影响力，产生良好的社会声誉，因此会受到对意见环境与普通用户情感的制约[41]。在衰退期，事件的相关讨论减少，公众的注意力被新的事件所分散，社交媒体中的海量信息和噪音阻断了社交机器人、意见领袖与普通用户间的情绪传导，用户情感表达受群体压力与意见环境的影响相对削弱，这使得三者情感的关联性减弱。这意味着如果不在事件初期对社交机器人进行识别和控制，社交机器人可能会影响到突发事件公众情感的整体走向。

5.2　社交机器人的影响策略具有隐匿性与间接性

在证实社交机器人情感影响的基础上，结合对社交机器人信息行为特征的描述，本研究发现了社交机器人影响策略的隐匿性与间接性特征。在隐匿性方面，社交机器人与普通用户缺乏互动，主要是作为微博转发者参与网络突发事件讨论；而且在本文选取的事件案例中社交机器人均无认证身身，较少在突发事件话题标签下参与讨论，微博内容较短，这决定了社交机器人发布的内容本身难以被其他用户发现和关注，因而社交机器人隐藏在普通用户群体中，游离在突发事件的讨论中心之外。社交机器人的间接性影响策略建立在其网络参与的隐匿性基础上。在微博社交媒体中，社交机器人并非通过直接与普通用户互动对普通用户施加影响，而是以一种间接方式影响网络环境：通过大量转发意见领袖的微博，增加意见领袖微博的转发量，提高意见领袖的微博在推荐系统中的排名，进而营造某种情感更占据主导的假象，强化意见领袖对普通网民的舆情信息影响。这支持了 M. Stella 等的研究结果，社交机器人具有“数字增强”的功能，能够提高个人或组织在在线活动中的影响力[42]。这种虚假人气会影响普通用户对事件的准确判断，制造虚假的认同。

同时，精细加工可能性模型(elaboration likelihood model, ELM)指出，信息接收者处理信息和态度改变的两条路径分别是中枢路径和边缘路径：通过中枢路径改变态度的信息接收者在知识、经验基础上对信息质量做出仔细评估，在边缘路径的信息处理中，态度改变主要是依靠与信息内容相关的情境因素[43]，如信源的粉丝数、社会资本等[44]。当突发事件发生时，社交机器人对意见领袖的转发行为提高了意见领袖在社交媒体中的可信性与话语权，这促使普通用户采取边缘路

径进行信息加工。而随着舆情演化,普通用户对信息的解释向中心路径转变,社交机器人和意见领袖对公众情感的影响呈现边际效应递减的特征。但是,社交机器人的参与仍然降低了微博的信息质量,破坏了社交媒体生态,为突发事件舆情埋下隐患。

本研究结果可为突发事件舆情治理提供参考。有关部门要建立和完善舆情监测系统,在事件爆发期,突发事件应急管理部门要关注网络意见领袖的参与行为,警惕参与用户激增现象与社交机器人对突发事件的非良性参与,及时发声稳定被煽动的网民情绪;在成熟期要梳理网民观点,与网民进行合理沟通,采取恰当方式回应网民诉求;在衰退期,要发挥舆论引导作用,维护事件的理性讨论,防止舆情反复。此外,社交机器人与意见领袖的密切联系为从源头治理社交机器人提供了可能。

6 结语

本研究以仁济医院赵晓菁事件为例分析社交机器人与意见领袖、普通用户情感在突发事件网络舆情不同阶段的情感交互关系。先前的研究集中于从网络关系角度研究社交机器人对人类的影响,忽视了宏观层面上社交机器人与其他主体情感关系的演化。为了实现这一点,我们在划分事件生命周期和提取微博文本情感的基础上对各阶段的社交机器人与意见领袖、普通用户的情感关系进行检验。本研究结果表明,社交机器人、意见领袖、普通用户的情感关系在不同阶段具有差异性,社交机器人主要在爆发期发挥影响,其影响策略具有隐匿性和间接性特点。该研究发现为将社交机器人嵌入突发事件舆情框架下的相关研究提供了参考。

本研究对网络突发事件舆情治理具有一定理论和实践意义,在理论上,本研究证实了网络突发事件中社交机器人、意见领袖与普通用户间存在情感关系,并进一步分析了情感关系在事件不同阶段的变化,丰富了网络突发事件舆情研究的主体和社交机器人研究的情境。在实践上,本研究发现的情感关系模式可为社交机器人识别提供参考;突发事件有关部门应该重视社交机器人的网络参与,依据各主体间情感关系的演化特征,在不同阶段采取相应措施,从而提高突发事件网络舆情治理效果。

同时,本研究也存在一些局限,本研究只在单个事件案例中对社交机器人与人类间的情感关系进行检验,社交机器人对人类情绪影响需考虑具体事件情境、机器人的数量以及网络的复杂程度。未来可以在更大规模的事件集上对社交机器人的情感关系进行验证。

参考文献:

[1] STIEGLITZ S, BUNKER D, MIRBABAIE M, et al. Sense-making in social media during extreme events[J]. Journal of contingencies and crisis management, 2018, 26(1): 4-15.
[2] HAGEN L, NEELY S, KELLER T E, et al. Rise of the machines? examining the influence of social bots on a political discussion network[J/OL]. Social science computer review, 2020. [2020-08-20]. https://journals.sagepub.com/doi/abs/10.1177/0894439320908190#articleCitationDownloadContainer. DOI: 10.1177/0894439320908190.
[3] ALLCOTT H, GENTZKOW M. Social media and fake news in the 2016 election[J]. Journal of economic perspectives, 2017, 31(2): 211-36.
[4] GUPTA A, LAMBA H, KUMARAGURU P, et al. Faking sandy: characterizing and identifying fake images on Twitter during hurricane sandy[C]//Proceedings of the 22nd international conference on World Wide Web. New York: ACM, 2013: 729-736.
[5] YANG K C, TORRES-LUGO C, MENCZER F. Prevalence of low-credibility information on twitter during the covid-19 outbreak[J]. arXiv preprint arXiv:2004.14484, 2020.
[6] 彭广林. 潜舆论·舆情主体·综合治理:网络舆情研究的情感社会学转向[J]. 湖南师范大学社会科学学报, 2020, 49(5): 142-149.
[7] FERRARA E, VAROL O, DAVIS C, et al. The rise of social bots[J]. Communications of the ACM, 2016, 59(7): 96-104.
[8] HOFEDITZ L, EHNIS C, BUNKER D, et al. Meaningful use of social bots? possible applications in crisis communication during disasters[C]//Proceedings of the 27th European conference on information systems. Stockholm & Uppsala, Sweden: AISeL. 2019: 1-16.
[9] BOSHMAF Y, LOGOTHETIS D, Siganos G, et al. Íntegro: leveraging victim prediction for robust fake account detection in large scale OSNS[J]. Computers & security, 2016, 61:142-168.
[10] HURTADO S, RAY P, MARCULESCU R. Bot detection in reddit political discussion[C]//Proceedings of the fourth international workshop on social sensing. New York: ACM, 2019: 30-35.
[11] DAVIS C A, VAROL O, FERRARA E, et al. Botornot: a system to evaluate social bots[C]//Proceedings of the 25th international conference companion on World Wide Web. New York: ACM, 2016: 273-274.
[12] KUDUGUNTA S, FERRARA E. Deep neural networks for bot detection[J]. Information sciences, 2018, 467:312-322.
[13] BESSI A, FERRARA E. Social bots distort the 2016 U.S. Presidential election online discussion[J]. First Monday, 2016, 21(11).
[14] BOICHAK O, JACKSON S, HEMSLEY J, et al. Automated diffu-

sion? bots and their influence during the 2016 US presidential election[C]//International conference on information. Cham: Springer, 2018: 17 - 26.

[15] BASTOS M T, MERCEA D. The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign[J]. Philosophical transactions of the royal society a mathematical physical and engineering sciences, 2018, 376(2128): 20180003.

[16] 师文,陈昌凤. 分布与互动模式:社交机器人操纵 Twitter 上的中国议题研究[J]. 国际新闻界,2020,42(5):61 - 80.

[17] 张洪忠,段泽宁,杨慧芸. 政治机器人在社交媒体空间的舆论干预分析[J]. 新闻界,2019(9):17 - 25.

[18] BRACHTEN F, STIEGLITZ S, HOFEDITZ L, et al. Strategies and influence of social bots in a 2017 German state election-a case study on Twitter[J]. arXiv preprint arXiv:1710.07562, 2017.

[19] YUAN X, SCHUCHARD R J, CROOKS A T. Examining emergent communities and social bots within the polarized online vaccination debate in Twitter[J]. Social Media + Society, 2019, 5(3): 2056305119865465.

[20] ALLEM J P, FERRARA E, UPPU S P, et al. E-cigarette surveillance with social media data: social bots, emerging topics, and trends[J]. JMIR public health and surveillance, 2017, 3(4): e98.

[21] BEIGI G, HU X, MACIEJEWSKI R, et al. An overview of sentiment analysis in social media and its applications in dis aster relief[M]//PEDRYCZ W, CHEN S M. Sentiment analysis and ontology engineering. Cham: Springer,2016: 313 - 340.

[22] 毛太田,蒋冠文,李勇,等. 新媒体时代下网络热点事件情感传播特征研究[J]. 情报科学,2019,37(4):29 - 35,96.

[23] STIEGLITZ S, DANG-XUAN L. Emotions and information diffusion in social media-sentiment of microblogs and sharing behavior[J]. Journal of management information systems, 2013, 29(4): 217 - 248.

[24] ZHANG W, WANG M, ZHU Y. Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? from the perspective of cognitive big data analytics[J]. International journal of information management, 2020, 50: 498 - 514.

[25] 周建,刘炎宝,刘佳佳. 情感分析研究的知识结构及热点前沿探析[J]. 情报学报,2020,39(1):111 - 124.

[26] ASIF M, ISHTIAQ A, AHMAD H, et al. Sentiment analysis of extremism in social media from textual information[J]. Telematics and informatics, 2020, 48: 101345.

[27] ASLAM F, AWAN T M, SYED J H, et al. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak[J]. Humanities and social sciences communications, 2020, 7(1): 1 - 9.

[28] 安璐,吴林. 融合主题与情感特征的突发事件微博舆情演化分析[J]. 图书情报工作,2017,61(15):120 - 129.

[29] 张鹏,崔彦琛,兰月新,等. 基于扎根理论与词典构建的微博突发事件情感分析与舆情引导策略[J]. 现代情报,2019,39(3):122 - 131,143.

[30] 邓君,孙绍丹,王阮,等. 基于 Word2Vec 和 SVM 的微博舆情情感演化分析[J]. 情报理论与实践,2020,43(8):112 - 119.

[31] 吴鹏,应杨,沈思. 基于双向长短期记忆模型的网民负面情感分类研究[J]. 情报学报,2018,37(8):845 - 853.

[32] LAZARSFELD P F, BERELSON B, GAUDET H. The people's choice[M]. New York: Columbia University Press, 1948:434 - 445.

[33] BAMAKAN S M H, NURGALIEV I, QU Q. Opinion leader detection: a methodological review[J]. Expert systems with applications, 2019, 115: 200 - 222.

[34] 王佳敏,吴鹏,陈芬,等. 突发事件中意见领袖的识别和影响力实证研究[J]. 情报学报,2016,35(2):169 - 176.

[35] 王国华,张剑,毕帅辉. 突发事件网络舆情演变中意见领袖研究——以药家鑫事件为例[J]. 情报杂志,2011,30(12):1 - 5.

[36] 李纲,陈璟浩. 突发公共事件网络舆情研究综述[J]. 图书情报知识,2014, 32(2):111 - 119.

[37] 谢科范,赵湜,陈刚,等. 网络舆情突发事件的生命周期原理及集群决策研究[J]. 武汉理工大学学报(社会科学版),2010,23(4):482 - 486.

[38] PESARAN H H, SHIN Y. Generalized impulse response analysis in linear multivariate models[J]. Economics letters, 1998, 58(1): 17 - 29.

[39] 高铁梅. 计量经济分析方法与建模[M]. 北京:清华大学出版社,2006:249,270.

[40] ZHANG L, WEI J, BONCELLA R J. Emotional communication analysis of emergency microblog based on the evolution life cycle of public opinion[J]. Information discovery and delivery, 2020, 48(3):151 - 163.

[41] 洪巍,王虎. 政府监管下的网络推手与意见领袖合谋行为研究——基于前景理论的演化博弈分析[J]. 经济与管理,2019, 33(3):18 - 25.

[42] STELLA M, CRISTOFORETTI M, DE DOMENICO M. Influence of augmented humans in online interactions during voting events[J]. PloS one, 2019, 14(5): e0214210.

[43] 张梦雅,王秀红. 精细加工可能性模型研究现状及应用领域分析[J]. 图书情报研究,2018,11(4):73 - 79,85.

[44] 王志英,邓航宇,王念新,等. 问答社区信息安全突发事件应急知识传播模型研究[J]. 情报杂志,2019,38(10):136 - 145.

作者贡献说明:

马晓悦:提供研究思路,设计研究方案,修订论文;
孟啸:设计方案,分析数据,撰写论文;
王镇:资料收集与整理;
刘益东:原始数据采集与处理。

Interactive Sentimental Influence Mechanism of Social Bots During an Emergency Event

Ma Xiaoyue[1]　Meng Xiao[1]　Wang Zhen[1]　Liu Yidong[2]
[1] School of Journalism and New Media, Xi'an Jiaotong University, Xi'an 710049
[2] Wei Re Dian Big Data Research Institute, Shanghai 201203

Abstract: [**Purpose/significance**] This study further elaborates the intervention mechanism of social bots on netizens' sentiment and provides suggestions for the management of public opinions in online emergencies. [**Method/process**] Taking the incident of Zhao Xiaojing in Renji Hospital as an example, with the Naive Bayes method calculating Weibo sentiment orientation, this study used the Granger causality test, impulse response analysis, and variance decomposition analysis based on the vector autoregression model (VAR) to judge the sentiment relationship between social bots, opinion leaders and ordinary users at all stages of the event life cycle. [**Result/conclusion**] The sentiment relationship between social bots, opinion leaders, and ordinary users changed with the evolution of the event. In the outbreak stage, social bots amplified the sentiment influence of opinion leaders on ordinary users. In the mature stage, the influence of social bots declined, and the sentiment of ordinary users reacted to social bots and opinion leaders. In the decline period, the three maintained a relatively independent sentiment relationship; Besides, the influence strategies of social bots were both insidious and indirect.

Keywords: emergency　social bots　sentiment analysis　time series analysis